SANDIA REPORT

SAND2014-0001
Unlimited Release
Printed November 2014

A New Perspective for the Calibration of Computational Predictor Models

Luis G. Crespo

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2014-0001
Unlimited Release
Printed November 2014

A New Perspective for the Calibration of Computational Predictor Models

Luis G. Crespo
V&V, UQ, and Credibility Processes
Sandia National Laboratories
P.O. Box 5800, Albuquerque, NM 87185-9999
lgcresp@sandia.gov

Abstract

This paper presents a framework for calibrating computational models using data from several and possibly dissimilar validation experiments. The offset between model predictions and observations, which might be caused by measurement noise, model-form uncertainty, and numerical error, drives the process by which uncertainty in the models parameters is characterized. The resulting description of uncertainty along with the computational model constitute a predictor model. Two types of predictor models are studied: Interval Predictor Models (IPMs) and Random Predictor Models (RPMs). IPMs use sets to characterize uncertainty, whereas RPMs use random vectors. The propagation of a set through a model makes the response an interval valued function of the state, whereas the propagation of a random vector yields a random process. Optimization-based strategies for calculating both types of predictor models are proposed. Whereas the formulations used to calculate IPMs target solutions leading to the interval value function of minimal spread containing all observations, those for RPMs seek to maximize the models' ability to reproduce the distribution of observations. Regarding RPMs, we choose a structure for the random vector (i.e., the assignment of probability to points in the parameter space) solely dependent on the prediction error. As such, the probabilistic description of uncertainty is not a subjective assignment of belief, nor is it expected to asymptotically converge to a fixed value, but instead it is a description of the model's ability to reproduce the experimental data. This framework enables evaluating the spread and distribution of the predicted response of target applications depending on the same parameters beyond the validation domain (i.e., roll-up and extrapolation).

Acknowledgment

This investigation was funded by the Verification, Validation, Uncertainty Quantification, and Credibility Processes Department of Sandia National Laboratories, Albuquerque, NM, USA.

Contents

List of Figures

Introduction

Model calibration refers to the process of prescribing the values of the parameters of a computational model according to experimental observations. Model-form uncertainty, measurement noise, and numerical error often prevent us from confidently prescribing a fixed constant value for such parameters. Consequently, a family of parameter values is prescribed such that the predictions resulting from evaluating the computational model at any member of this family are a sufficiently accurate representation of the observations [4].

Several model calibration techniques are available in the literature. Many of them assume the structure

$$y = M(x, p) + \eta, \tag{1}$$

where $y \in \mathbb{R}$ is the *response* or model's output, M is the computational model, $x \in \mathbb{R}^n$ is the *state* or explanatory variable, $p \in \mathbb{R}^{n_p}$ is the model's parameter, and $\eta \in \mathbb{R}$ is a random variation caused by noise and measurement error. Many of the model calibration techniques are based on this model structure, the assumption of p being a fixed but unknown constant (i.e., the uncertainty in p is epistemic), and the assumption of η being independent and identically distributed (IID) following a normal distribution with zero mean and a fixed variance. In contrast, the framework proposed herein is applicable to models M, where p can either be epistemic or aleatory (i.e., uncertainty caused by inherent, and irreducible variability) taking on an arbitrary form[1].

A typical regression problem consists of estimating the value of p in M given the set of observations (x_i, y_i), $i = 1, \ldots, N$, where $N > n_p$. This is often carried out by searching for the parameter realization that minimizes the sum of squared residuals, p_{LS}. The precision of this estimate, which prescribes how much it can deviate from its "true value" within an epistemic framework, is often evaluated using confidence intervals [8, 5]. The calculation of confidence intervals, prediction intervals (i.e., intervals where future observations are expected to fall) and credible intervals [4] require having a probabilistic description of the uncertainty in p. This description often requires (i) knowing the distribution for the prediction errors, (ii) M and η taking particular forms (e.g., M depends linearly on p and the noise η is additive) or (iii) M being accurately represented by a linear approximation. As such, the suitability of the resulting confidence and prediction intervals depends tigthly on the validity and accuracy of the underlying assumptions and approximations.

The most common approach to model calibration is Bayesian inference. In parametric Bayesian inference the objective is to describe the model's parameters as a vector of possibly dependent random variables using Bayes' rule [4]. The resulting vector, called the posterior, depends on an assumed prior random vector, and the likelihood function; which in turn depends on the set of N observations available and on the model structure of M. Whereas this approach does not make any limiting assumptions on the manner in which M depends on p, nor on the structure of the resulting posterior; it requires for p to be epistemic. The

[1] Equation (1) can be seen as a particular case, in which all epistemic uncertainties enter into M, whereas η is a single additive aleatory uncertainty having a known structure.

presence of aleatory uncertainty and model-form uncertainty yields uncertainty characterizations for p that fail to describe the prediction error. This can be mitigated by adding a non-deterministic discrepancy term to M [4]. This practice however, prevents rolling-up the resulting prescription of uncertainty through other computational models. By roll-up we mean the process of calibrating the parameters of one or several computational models and using the resulting characterization of p to make predictions of a target application which also depends on p. In spite of this limitation, its high computational demands, and of the potentially high sensitivity of the posterior to the assumed prior, this method is commonly regarded as the benchmark for model calibration.

This paper presents techniques for characterizing the uncertainty in parameters common to several computational models based on multiple validation experiments. These experiments might measure dissimilar quantities of interest having different state variables ranging over different domains (e.g., some might be temperatures at particular locations whereas others might be heat flux at other locations). For each of these validation experiments we have a computational model, and a set of observations. Uncertainty is characterized according to the dispersion of the prediction errors. This characterization and a computational model constitute a predictor model. This paper focuses on Interval Predictor Models (IPM), for which uncertainty is characterized as a hyper-rectangular set[2], and Random Predictor Models (RPM), for which uncertainty is characterized as a random vector supported in a hyper-rectangular set. Hyper-rectangular sets are preferred since they have the advantage that the range of an individual parameter is not affected by the value taken by the others. IPMs and RPMs are used to generate informative predictions of the spread of the response at values of the state that might extend beyond the range of observations (i.e., extrapolation). Such predictions are based on the intrinsic structure of the computational model and on its ability to describe the observations.

The calibration of models that depend linearly in p and polynomial in x can be done rigorously [3]. This process yields a formal description of the *model's reliability*, i.e., the probability that future observations will lie within the predicted range of responses by analytical, non-statistical means, of the uncertainty in p, and of the spread in the corresponding response y. This paper extends these ideas to cases in which (i) data from multiple validation experiments are available (i.e., $n_v > 1$, where n_v is the number of validation experiments), (ii) the model structure depends arbitrarily on parameters and states and might not even have an explicit form (e.g., M is a finite element model predicting the stress at a particular location of a nonlinear structure), and (iii) the resulting characterization of uncertainty is probabilistic. This generality precludes the formal descriptions listed above.

This paper is organized as follows. Section introduces basic notions and definitions and describes the main objectives of this article. This is followed by Section where a framework for calculating IPMs is presented. Section introduces RPMs and a formulation for their calculation. This is followed by Section where the methodology is extended to the roll-up problem. Finally, some conclusions are made.

[2]By hyper-rectangular set, we mean a set which can be represented by a Cartesian product of intervals. This constrains its orientation so that its edges are parallel to the coordinate axes.

Problem Statement

In the developments that follow we assume that experimental observations are stationary and independent, that the response of the models depend continuously on both the parameters and the states, and that the models have a deterministic mathematical structure, i.e., they yield a fixed value of the response for fixed values of the state x and of the parameter p. Consider the ensemble[3] of computational models $y^j = M^j(x^j, p)$, for $j = 1, \dots n_v$, where the response y^j is a scalar, the state x^j is a vector of length n_j, $p \in \mathbb{R}^{n_p}$ is a parameter, and n_v is the number of validation experiments. Note that all n_v computational models share the same parameter p. The values of parameters that are not common to all n_v models can be set to fixed constant values, e.g., the corresponding term in the least squares solution, before applying this methodology. This practice enables rolling up the uncertainty in all calibrated models to system-level target applications.

Further assume we have experimental observations corresponding to each of the n_v validation experiments. Denote by N^j the number of observations corresponding to the jth validation experiment. The Data-Generating Mechanism (DGM) associated with each of these experiments is postulated to act on the state to produce a response. The responses can depend on state variables and on some other influences such as intrinsic variability and measurement noise. Let $X^j \subseteq \mathbb{R}^{n^j}$ be a set of state variables, and $Y^j \subseteq \mathbb{R}$ be the set of responses corresponding to the DGM governing the jth validation experiment acting on elements of X^j. The data sequence associated with this experiment is denoted as $\mathsf{z}^j = (x_i^j, y_i^j)$, where $i = 1, \dots N^j$. The DGMs might depend on parameters and states that are absent from the computational models. As compared to a DGM, the model M^j might be subject to parametric and model-form uncertainty. Parametric uncertainty can be epistemic, aleatory or a combination of both. As compared to a computational model with the highest fidelity possible, M^j often suffers from approximation, numerical, and discretization errors. These errors are a consequence of making M^j computationally tractable.

The first objective of this paper is to prescribe p according to the data sequences z^1, z^2, ..., z^{n_v}, based on the predictive ability of the computational models M^1, M^2 ..., M^{n_v} describing such experiments. The second objective is to use this prescription to generate informative predictions corresponding to unobserved realizations of the state. An informative prediction can be interpreted as a prediction that is consistent with salient features of the data sequences and with the predictive ability of the n_v computational models used to characterize p. In the context of IPMs, a prediction is an interval of possible responses whereas for RPMs a prediction is a random variable. The spread and probability concentration of such predictions are driven by the model's ability to reproduce the observations. As such, models that fail to properly capture the physics driving the DGM will yield large uncertainties and wide prediction ranges.

[3]The following conventions will be used regarding the usage of subindices and superindices on the model's state and model's response. Subindices will be used to describe experimental realizations whereas superindices will be used to describe different states/responses; e.g., y_3^2 is the third observed response associated to model M^2 whose state and response are x^2 and y^2 respectively.

For simplicity in the presentation, the two sections on IPMs and RPMs that follow assume that there is only one validation experiment (i.e., $n_v = 1$). Consequently, we will drop the super-indices from the notation of states and responses.

Interval Predictor Models

An IPM is simply a function that returns an interval as output. These models were proposed in the framework of differential inclusions and set-valued dynamical systems [1, 2], and further developed for the case in which prior knowledge on the data-generating mechanism is available [6, 7]. In the context of this paper, an IPM is a rule that assigns to each instance vector $x \in X$ a corresponding outcome interval in Y. That is, an IPM is a set-valued map

$$I : x \to I_y(x) \subseteq Y, \tag{2}$$

where x is a state on which the model's output depends, and $I_y(x)$ is the prediction interval. Let M be any functional acting on a vector x of state variables and a vector p of parameters to produce an output y; i.e., $y = M(x, p)$. A parametric IPM is obtained by associating to each $x \in X$ the set of all possible outputs y that result from varying p over some subset P of parameter space:

$$I_y(x, P) = \{y : y = M(x, p) \text{ for all } p \in P\}. \tag{3}$$

$I_y(x, P)$ will be an interval as long as $M(x, p)$ is a continuous function of x and p, and P is a connected set. All instances of M and P considered in this paper will satisfy these restrictions. Attention will be limited to the case in which P is the bounded[4] hyper-rectangle: $P = \{p : p_{\min} \leq p \leq p_{\max}\}$. The prediction of the IPM can be described as $I_y(x, P) = [\,\underline{y}(x, P),\ \overline{y}(x, P)\,]$, where

$$\underline{y}(x, P) = \min_{p \in P}\{M(x, p)\}, \tag{4}$$

$$\overline{y}(x, P) = \max_{p \in P}\{M(x, p)\}. \tag{5}$$

The functions $\underline{y}$ and $\overline{y}$ are envelopes of the family of infinitely many predictions that result from evaluating $M(x, p)$ for each parameter realization $p \in P$. Note that the envelopes themselves might not be members of the family (e.g., the envelopes of a family of polynomials need not to be polynomials). The spread of I_y, which is the separation between the envelopes, is

$$\delta_y(x, P) = \overline{y}(x, P) - \underline{y}(x, P). \tag{6}$$

The solutions to Equations (4-5) occur at vertices of P when M is a linear function of p (the particular vertex might vary with x). Otherwise, the solution might correspond to an interior point of P. The case in which M is a convex function of p can be efficiently solved by using convex optimization techniques (when solving for $\underline{y}$) and by evaluating M at the

[4]Vector inequalities hold component-wise.

vertices's of P (when solving for $\overline{y}$). If none of the above cases hold, nonlinear optimization or sampling techniques can be readily used.

The DGM is often approximated by the Least Squares (LS) prediction $y = M(x, p_{\text{LS}})$, where the parameter estimate p_{LS} is given by

$$p_{\text{LS}} = \underset{p}{\text{argmin}} \left\{ \sum_{i=1}^{N} \left(y_i - M(x_i, p) \right)^2 \right\}. \tag{7}$$

Whereas the LS prediction describes the overall trend of the data as described by a point estimate of p, $I_y(x)$ describes its spread. A set P, which prescribe such a spread, might not even contain p_{LS}.

Formulations for calculating two types of IPMs are presented next. These formulations parameterize the geometry of P differently. Denote by θ the set of parameters required to fully prescribe P. The two parameterizations to be considered are (i) $\theta = \{c, m, \alpha\}$, where c is the geometric center, m is the unit diagonal vector, and $\alpha > 0 \in \mathbb{R}$ is the length of the semi-diagonal of P; and (ii) $\theta = \{p_{\min}, p_{\max}\}$, where $p_{\min}$ and $p_{\max}$ are the "lower left" and "upper right" corners of P. The dependence among these variables is given by

$$c = \frac{p_{\max} + p_{\min}}{2}, \;\; m = \frac{p_{\max} - p_{\min}}{\|p_{\max} - p_{\min}\|_2}, \;\; \alpha = \frac{\|p_{\max} - p_{\min}\|_2}{2}. \tag{8}$$

Hyper-rectangular Sets with Fixed Parameters

Here we seek the IPM given by (3) where the uncertainty set $P(\theta)$ is parameterized by $\theta = \{c, m, \alpha\}$ according to

$$P(\theta) = \{p : c - \alpha m \;\leq\; p \;\leq\; c + \alpha m\}. \tag{9}$$

The following optimization program yields an IPM whose uncertainty set (9) has fixed given values for c and m whereas α is optimized.

The Optimization Programs in this paper utilize $E_x[\cdot]$, the expected value operator with respect to x, in their objective functions. When x has a known distribution, the implied integral can be evaluated analytically or numerically depending on the integrand. Otherwise, the sample mean of the integrand using the data in z can be used to approximate it.

Optimization Program 1. *Consider the IPM defined by (3) and (9). Assume that the values c and $m > 0$ are set in advance and denote by $E_x[\cdot]$ the expectation operator with respect to the state variable x. The value of α required to fully prescribe this IPM is*

$$\hat{\alpha} = \underset{\alpha > 0}{\text{argmin}} \left\{ E_x[\delta_y(x, P)] : \underline{y}(x_i, P) \leq y_i \leq \overline{y}(x_i, P),\; 1 \leq i \leq N \right\}. \tag{10}$$

Therefore we search for the set P with a fixed geometric center and unit diagonal vector that minimizes the expected interval spread such that all the observed responses are within

the limits of the interval valued function I_y. The corresponding uncertainty set, denoted as $\hat{P}$, is given by (9) evaluated at $\hat{\theta} = \{c, m, \hat{\alpha}\}$. Note that if $P_1 \subset P_2$, then $\delta_y(x, P_1) \subseteq \delta_y(x, P_2)$, and $E_x[\delta_y(x, P_1)] \leq E_x[\delta_y(x, P_2)]$. Consequently, the minimization in (10) yields uncertainty sets of minimal size in this sense.

A few comments regarding the selection of the fixed parameters in P are in order. A natural candidate for c is p_{LS}. Even though this practice leads to a prescription of uncertainty in which the prediction corresponding to the center of the box optimally describes the overall trend of the data, the corresponding spread in the prediction might be suboptimal (i.e., the selection of another center point leads to a smaller expected spread). Regarding the unit diagonal vector, the components of m should be made proportional to the expected range of uncertainty in the corresponding parameters. Note that m should not only account for such ranges but also for the different units in the corresponding p's.

The above formulation does not take into account uncertainty in experimental observations. Error due to instrument calibration, data acquisition, and variation in experimental conditions yield random and bias errors. This uncertainty is often prescribed as error bars centered about the measurements, e.g., the uncertainty in (x_i, y_i) is cast as $(x_i \pm dx, y_i \pm dy)$ where $dx > 0$ and $dy > 0$. The effects of experimental uncertainty can be taken into account by replacing the inequality constraints in (10) with

$$\underline{y}(x, P) - dy \leq y_i \leq \overline{y}(x, P) + dy, \text{ for all } x \in [x_i - dx,\ x_i + dx]. \tag{11}$$

The particular case in which $dx = 0$ yields $\overline{y}(x_i, P) + dy \geq y_i \geq \underline{y}(x_i, P) - dy$. The general form of (11) is hard to evaluate since the inequality applies to a continuum of state values. A natural approximation to (11) is given by $\underline{y}(x_k^v, P) \leq y_k^v \leq \overline{y}(x_k^v, P)$ for $k = 1, \ldots, 2^{n^j+1}$, where $\{(x_k^v, y_k^v) : 1 \leq k \leq 2^{n+1}\}$ enumerates the vertices of the box $[x_i - dx,\ x_i + dx] \times [y_i - dy,\ y_i + dy]$. The formulations herein ignore experimental uncertainty, thus $dx = 0$ and $dy = 0$. Their extension to the general case can be carried out by taking these considerations into account.

The formulation in (10) as well as some of the concepts to be introduced later in the paper can be described using the *Zero-Prediction-Error Manifold*. This is the manifold of parameter points for which the computational model M can exactly reproduce an observation point. More specifically, the Zero-Prediction-Error Manifold for the data point (x_i, y_i) is defined as

$$S_i = \{p : y_i - M(x_i, p) = 0\}. \tag{12}$$

Note that a feasible solution to (10) leads to a set P whose intersection with S_1, S_2, ..., S_N are all non-empty. When M depends linearly on p such manifolds are hyperplanes. When the response y_i is outside the range of predictions associated with the model at x_i, S_i is empty. When the observation is within such a range, there might be infinitely many parameter realizations and, thus, infinitely many fixed parameter point predictions $M(x_i, p)$ passing through y_i

The separation between c and the manifold S_i in (12) can be evaluated via the *Parametric*

Safety Margin (PSM), which is defined as

$$\rho_i = \min_p \{\|c - p\|_m^\infty : y_i = M(x_i, p)\}, \tag{13}$$

where the m-infinity norm of a vector a is defined by

$$\|a\|_m^\infty = \max_{1 \le k \le n_p} \left\{ \frac{|a_k|}{m_k} \right\}.$$

The PSM ρ_i is the length of the semi-diagonal of the smallest hyper-rectangle oriented as $\hat{P}$ (i.e., same center and same diagonal orientation) for which one of the corresponding IPM envelopes passes through the ith observation point. Hence, ρ_i evaluates the extent by which the presence of the data point (x_i, y_i) contributes to the spread of $\hat{P}$. The spread of the resulting prediction and of the uncertainty set are driven by the observation(s) attaining the largest PSM values. Note however, that the particular observations in z attaining such values vary with c and m. Equation (13) is equivalent to

$$\rho_i = \min_{p,\alpha} \{\alpha : \alpha > 0,\ y_i = M(p, x_i),\ -\alpha m \le c - p \le \alpha m\}, \tag{14}$$

which, as opposed to (13), does not suffer from derivative discontinuities. Note that $\hat{\alpha} = \max_{i=1\ldots N}\{\rho_i\}$. PSMs are used in [3] to identify and eliminate outliers from the data set and to calculate IPMs having tighter predictions.

Fixed-parameters IPM Example

In this example we will consider *Validation Experiment 1*. This experiment consists of $N^1 = 50$ observations of an output which depends on a single state variable ($n_x^1 = 1$) and two parameters ($n_p = 1$). This DGM depends nonlinearly on the state and parameters, and it is subject to state-dependent measurement noise. Figure 1 shows the elements of the corresponding data sequence, z^1.

A mathematical model $M^1(x^1, p)$ describing this DGM was built. This model, which depends nonlinearly on $x^1 \in \mathbb{R}$ and $p \in \mathbb{R}^2$, suffers from model-form and epistemic uncertainty. The prediction corresponding to the LS parameter, which is $p_{\text{LS1}} = [0.9420, 1.0689]^\top$, and the prediction interval function [5] associated with the nonlinear regression are also shown in Figure 1. In statistical inference [8], a prediction interval is an estimate of an interval in which future observations will fall, with a certain probability, given what has already been observed. Note that more than 60% of the observations lay outside the interval valued function.

The model $M^1(x^1, p)$ and the $N^1 = 50$ observations were used to build an IPM. The resulting predictor model will be called IPM1. The solution to (10) yields $\hat{\alpha} = 0.2625$, and $E_x[\delta_y] = 2.2505$. Figure 2 shows the envelopes of IPM1 for $c = p_{\text{LS}}$ and $m = [\sqrt{0.5}, \sqrt{0.5}]^\top$. All 50 observations are within the envelopes and the lower envelope passes through one of them, indicating that the corresponding constraint is binding. This particular observation

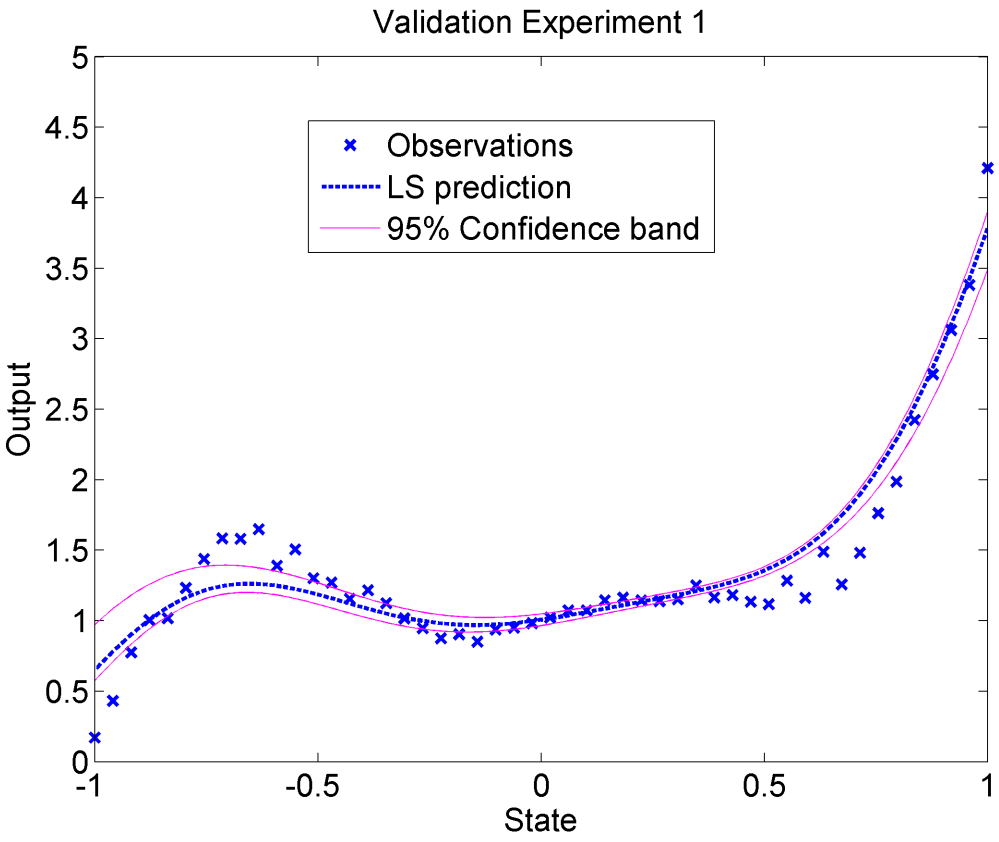

Figure 1. Observations, LS prediction, and confidence band.

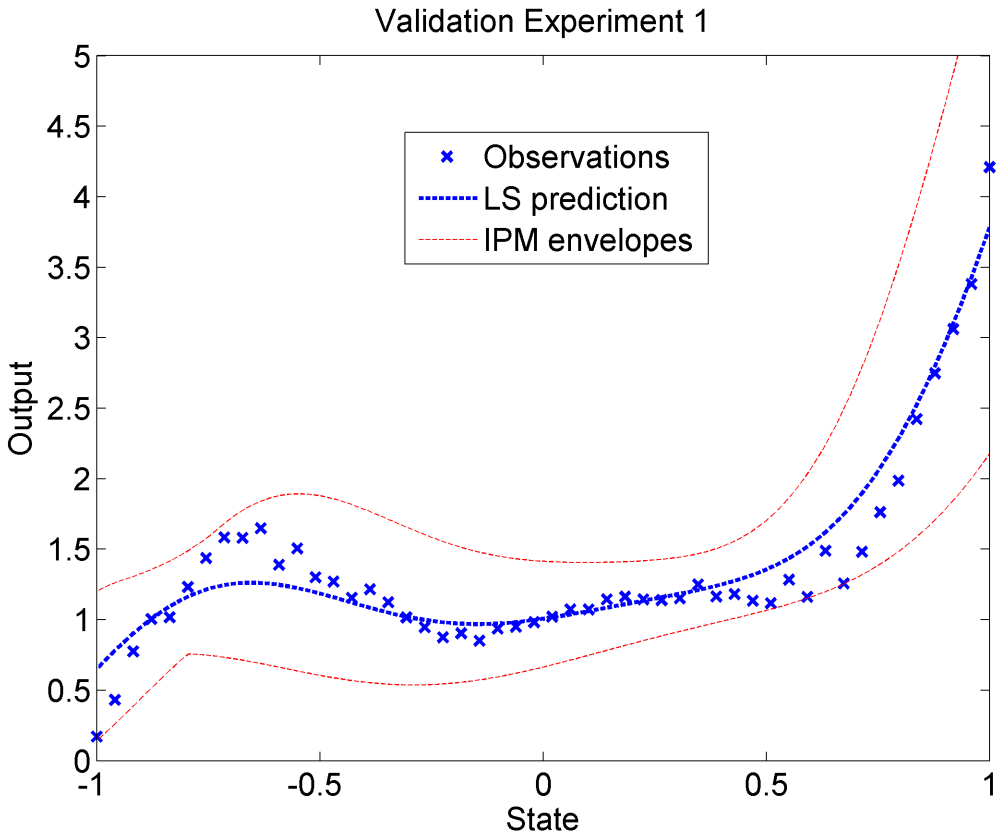

Figure 2. IPM with fixed parameters.

point, located near $(0.7, 1.25)$, ends up prescribing the extension of $\hat{P}$. Recall that the functional form of the envelopes depends exclusively on the observations and the structure of M^1. Whereas $M^1(x^1, p_{\text{LS}})$ describes the dominant trend of the data by weighing *all* data points equally, the interval valued function $I_{y^1}(x^1, \hat{P})$ describes their spread. This spread, which is driven by extreme observations, is not necessarily centered about $M^1(x^1, p_{\text{LS}})$. In general, there is no basis to expect that (i) $p_{\text{LS}} \in \hat{P}$, or that (ii) $\underline{y}^1(x^1) \leq M^1(x^1, p_{\text{LS}}) \leq \overline{y}^1(x^1)$ for all $x^1 \in X$, or that (iii) the centerline of $I_{y^1}(x^1, \hat{P})$, $(\underline{y}^1(x^1, \hat{P}) + \overline{y}^1(x^1, \hat{P}))/2$, accurately describes the overall trend of the data.

Hyper-rectangular Sets with Free Parameters

In this formulation we seek the IPM in (3) for the uncertainty set $P(\theta)$ parameterized by $\theta = \{p_{\min}, p_{\max}\}$ according to

$$P(\theta) = \{p : p_{\min} \leq p \leq p_{\max}\}, \tag{15}$$

such that all components of θ are optimal.

Optimization Program 2. *Consider the IPM defined by (3) and (15). This IPM is fully prescribed by the limits of P, which are given by*

$$\langle \hat{p}_{min}, \hat{p}_{max} \rangle = \underset{p_a,\, p_b}{\operatorname{argmin}} \{ E_x[\delta_y(x, Q)] : \ \underline{y}(x_i, Q) \leq y_i \leq \overline{y}(x_i, Q),\ 1 \leq i \leq N, \qquad (16)$$
$$Q = P(\theta) \text{ where } \theta = \{p_a, p_b\},\ p_a \leq p_b \}.$$

Therefore we search for the limits of P that minimize the expected interval spread such that all the observed responses are within the limits of the interval valued function $I_y(x)$. The resulting uncertainty set, $\hat{P}$, is given by (15) evaluated at $\hat{\theta} = \{\hat{p}_{\max}, \hat{p}_{\min}\}$. Remarks regarding the evaluation of $E_x[\delta]$ made in the previous section apply here as well. The PSM associated to resulting IPM are given by (13) with the parameters c and m corresponding to $\hat{P}$.

The additional flexibility of (16), as compared to (10), often yields IPMs with tighter predictions and smaller uncertainty sets. As before, the spread of both the prediction and the uncertainty set are often driven by a few data points. These data points, which attain the largest PSM values from the full ensemble in z, often deviate significantly from the rest of the observations. The empirical Cumulative Distribution Function (CDF) of the PSMs is defined to be the function whose value at an arbitrary real number ρ is the fraction of the N values of the PSMs which are less than or equal to ρ. Examining this empirical CDF enables visualizing how likely are the extreme observations prescribing the IPM envelopes. The majority of PSM values might be considerably smaller than the largest PSM value(s) of the set, even though the IPM is prescribed according to the data point(s) close to such values.

Free-parameters IPM Example

The free-parameter formulation in (16) was used to build an IPM based on the same computational model and the same observations used in the previously. The resulting model, to be referred to as IPM2, is shown in Figure 3. The optimal solution leads to $c = [0.8837, 0.9740]^\top$, $m = [0.3236, 0.9462]^\top$, and $E_x[\delta_y] = 1.9642$. The improved tightness in the model prediction of IPM2 as compared to that of IPM1 is apparent. Note that most improvements are reflected in a tighter upper envelope, while the lower envelope degrades slightly. Further notice that the LS prediction is closer to upper envelope than in the IPM1 case.

Figure 4 shows the empirical CDF of the PSM associated with IPM2. The vertical lines are the 10-percentiles. Note that the PSM values range from 0.0005 to 0.3263, but about 80% of them take on values that are less than 0.15. The observation(s) for which the PSM attains the largest value(s) prescribe the IPM whereas the rest of them are inconsequential. The optimization programs above can be cast as min-max problems since they lower the worst-case PSM (i.e., the 100-percentile of ρ) as much as possible. When outliers are present in the data set, or when the analyst wants tighter predictions for most of the observations (i.e., lower values of ρ), the minimization of lower percentiles can be considered instead [3].

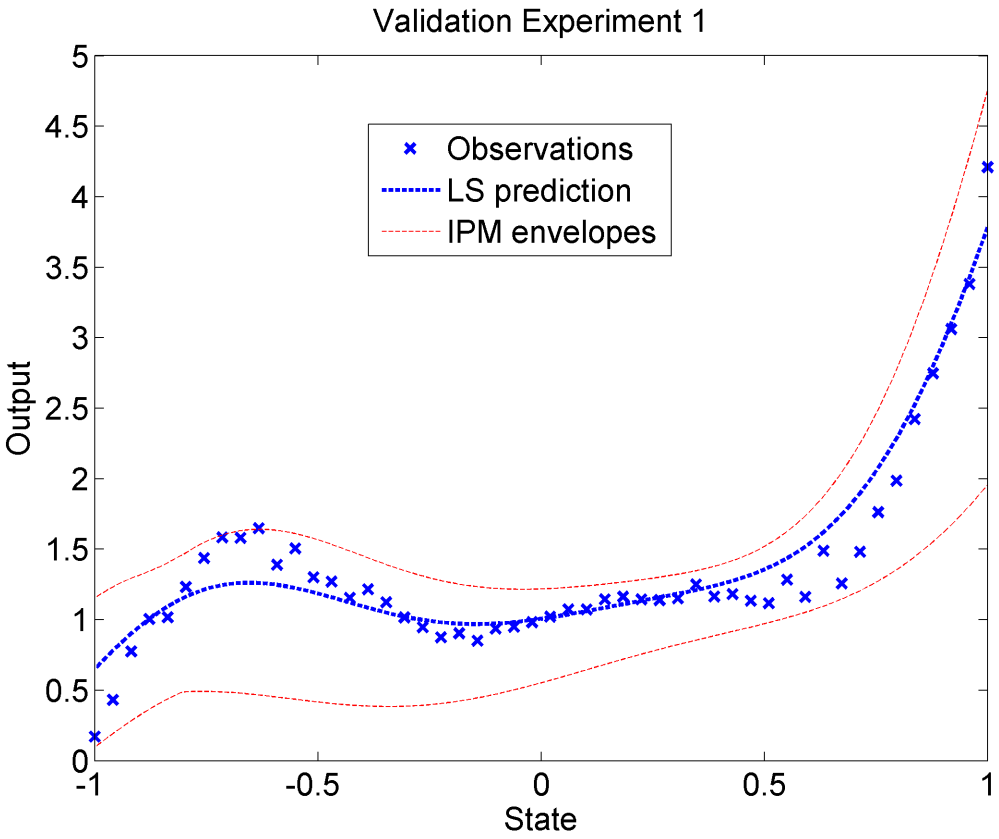

Figure 3. IPM with free parameters.

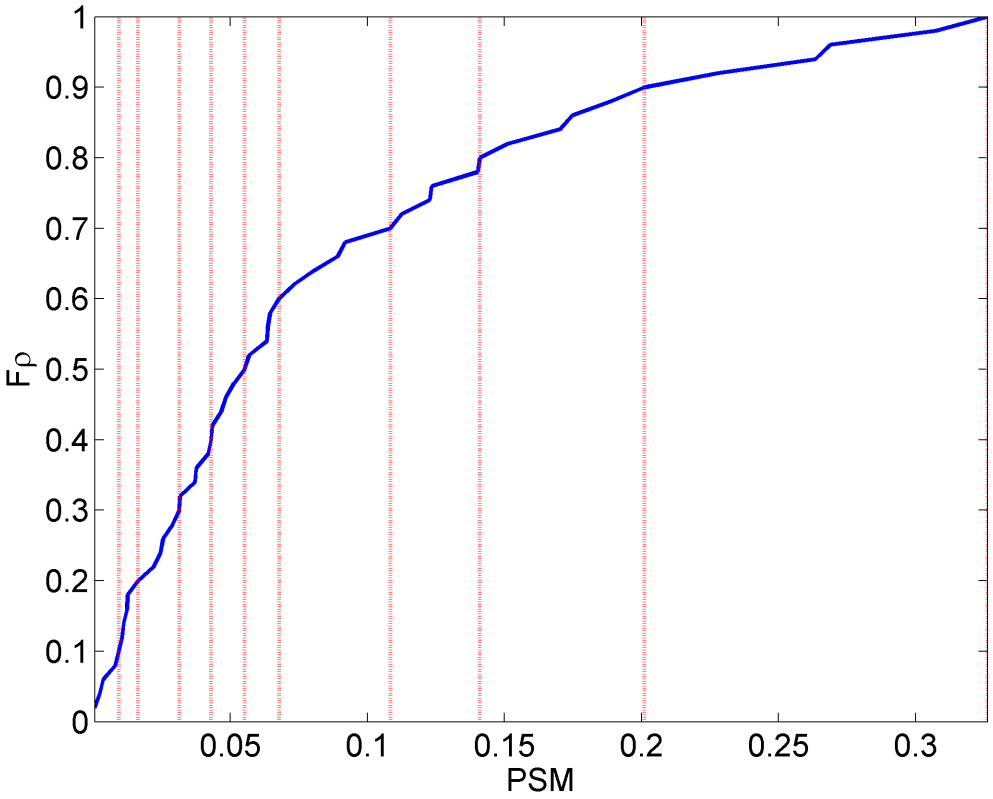

Figure 4. Empirical CDF of the PSMs associated to IPM2

IPMs are effective for describing extreme and possibly unlikely predictions but lack the fidelity to describe the likelihood of occurrence within $I_{y^1}(x^1)$. RPMs, to be introduced next, provide a means to mend for this deficiency.

Random Predictor Models

An RPM is simply a rule that assigns to each state vector $x \in X$ a corresponding random variable in the output space Y. That is, an RPM is a random variable-valued map

$$R : x \to R_y(x) \subseteq Y, \tag{17}$$

where x is the state, and $R_y(x)$ is a random process having a support that lies in Y. The prediction of the RPM is

$$R_y(x) = \{y : y = M(x,p),\ p \sim f_p(p) \text{ for } p \in P(\theta)\}, \tag{18}$$

where $f_p(p)$ is the joint Probability Density Function PDF of p having $P(\theta)$ as its support set. As with the IPMs presented above, we will focus on hyper-rectangular P's parameterized by θ. When both P and M are bounded, which is the case assumed in this paper, the support of $R_y(x)$ is bounded as well. The PDF of the predicted response $M(x,p)$ at any value of x is fully prescribed by that of p. The PDF of the random variable $R_y(\hat{x})$ will be denoted as $f_{M(\hat{x},p(\theta))}(y)$. Therefore, statistics of the response, such as the mean $\mu_y(x) = E_p[M(x,p)]$, the variance $\nu_y(x) = E_p[(M(x,p) - \mu_y(x))^2]$, and the support $I_y(x)$ (given in Equations (4,5)); vary with x. Note that the limits of the support of an RPM are the envelopes of an IPM having the same P.

Figure 9 shows an example of the RPM in (18). This figure shows the LS prediction, the expected/mean response, and the 5-percentile curves for a uniform PDF supported in the $\hat{P}$ corresponding to IPM2. This RPM, which will be compared to other RPMs below, will be referred to as RPM1. Recall that IPM2 leads to a range of predictions with minimal spread. Note the sizable offset between the LS prediction and the expected response $\mu_y(x)$. It turns out that the mean response underestimates 41 of the 50 experimental observations by a considerable margin, thus, it is a poor descriptor of the overall trend of the data. Further notice that the probability of occurrence of the data sequence z under the IID assumption, which is the product of the individual probabilities at the observation points and whose values can be inferred from the percentile curves, is comparatively low (see Figure 8 for comparison).

This paper proposes a process for constructing a probabilistic description for p, $f_p(p)$ with $p \in P(\theta)$, based on a non-subjective figure of merit. This figure of merit rewards parameter realizations (i.e., assigns a higher value) leading to model predictions that closely describe the experimental observations (i.e., points in the parameter space lying on the manifolds S in (12)). The closer the parameter point to the manifolds associated to all the experimental observations, the larger the value of the figure of merit, thus, the larger the value of the joint

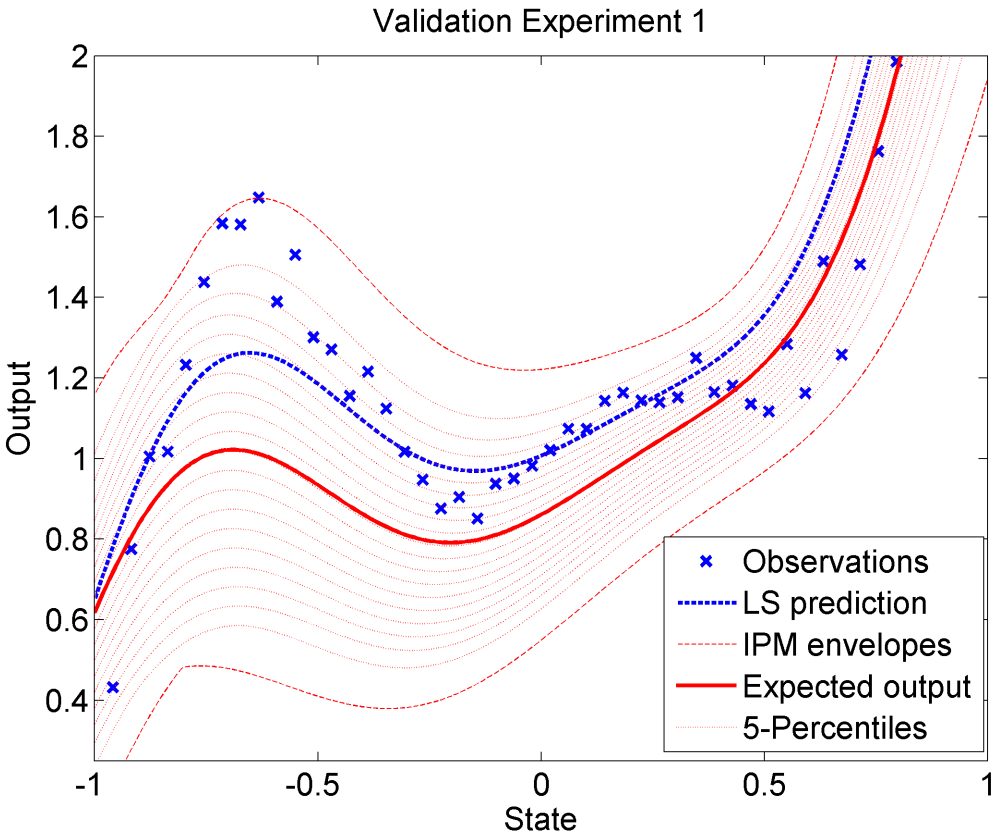

Figure 5. RPM for a uniform joint PDF (RPM1).

density function at such a point. Note that the "distance" from any parameter point to any of such manifolds, which can be evaluated using several different norms, will only depend on the experimental observation (x_i, y_i) and on the ability of the computational model $y = M(x, p)$ to replicate such an observation. Note that these two underlying criteria also prescribe the likelihood function used in Bayesian inference. The developments that follow propose a few norms to evaluate such a "distance", thus, a figure of merit, and describe a procedure from which probabilistic descriptions of p, consistent with this figure of merit, can be built.

Assume the joint PDF for p is given by

$$f_p(p) = \frac{\gamma(p)}{C_0} \text{ for } p \in P(\theta), \tag{19}$$

where C_0 is a normalization constant so the volume under the joint PDF over $P(\theta)$ is one, and θ is a variable to be determined. The function $\gamma(p)$ assigns a figure of merit to each possible realization of p. Denote by $d(p, S_i) \geq 0$ a norm of the separation between the parameter point p and the zero prediction error manifold in (12). The structure of the function γ to be considered is given by

$$\gamma(p) = \left(\beta_1 + \beta_2 \sum_{i=1}^{N} d(p, S_i) \right)^{-1}, \tag{20}$$

where $0 < \beta_1 \ll 1$ is a small number used to prevent γ from being unbounded, and $\beta_2 > 0$ is a scaling parameter used to control the smoothness of $\gamma(p)$. The closer p is to parameter points where the computational model is able to reproduce all experimental observations, the larger the value of γ, thus, the larger the value of $f_p(p)$ at p. Note that in an idealized case in which there is no measurement noise or model-form uncertainty, and the only source of uncertainty is epistemic, $\gamma(p)$ will be maximal at the "true" value(s) of p.

Several norms $d(p, S_i)$ can be used to fully prescribe (20). Each of them will lead to a different joint PDF, thus to a different random process $R_y(x)$. Three suitable norms are

$$d(p, S_i) = \min_{\bar{p}} \{ \|p - \bar{p}\|_2 : y_i = M(x_i, \bar{p}) \}, \tag{21}$$

$$d(p, S_i, m) = \rho_i = \min_{\bar{p}} \{ \|p - \bar{p}\|_m^\infty : y_i = M(x_i, \bar{p}) \}, \tag{22}$$

and

$$d(p, S_i) = \frac{(y_i - M(p, x_i))^2}{C_i}, \tag{23}$$

where C_i is a normalization constant used to ensure that $d(p, S_i) \leq 1$ for all $i = 1 \ldots N$. The norms in (21) and (22) measure the separation between any parameter point p and the zero-prediction-error manifold by evaluating a distance in the parameter space. Conversely, (23) measures separation by evaluating a distance in the output space. As such, the norm in (23) yields results that are structurally different from those based on the other two norms (unless M is linear in p). Furthermore, the norms in (21) and (22) require solving an optimization problem, whereas (23) only requires a single model evaluation. This makes the norm in (23) preferable from a computational stand point. The particular case in which M depends

linearly on p and polynomially on x, i.e., $y = p^\top \varphi(x)$ where $\varphi(x)$ is a vector of monomials in the components of x, enables evaluating (21) and (22) analytically. In such a case, we obtain

$$d(p, S_i) = \frac{|y_i - p^\top \varphi(x_i)|}{\|p\|_2}, \tag{24}$$

$$d(p, S_i, m) = \left| \frac{y_i - \varphi(x_i)^\top p}{\varphi(|x_i|)^\top m} \right|. \tag{25}$$

Once the norm d, and the value of the parameter θ have been chosen, the uncertainty model $f_p(p)$, thus, the RPM in (18) are fully prescribed.

The formulation below targets RPM with optimal characteristics. Two performance metrics used to quantify desirable traits of the prediction will be considered. One of them is

$$\begin{aligned} J(\theta, f_p, \mathsf{z}) &= \frac{1}{N} \sum_{i=1}^{N} \int_{p_{\min}}^{p_{\max}} f_p(p, \theta)(M(x_i, p) - y_i)^2 dp, \\ &= \frac{1}{N} \sum_{i=1}^{N} \left(V[M(x_i, p)] + (E[M(x_i, p)] - y_i)^2 \right), \end{aligned} \tag{26}$$

where $E[\cdot]$ and $V[\cdot]$ are the expectation and variance operators. This metric, which evaluates the dispersion of the random process $R_y(x)$ about the observations, is computationally cheap to evaluate since it only requires evaluating means and variances. Another metric is

$$J(\theta, f_p, \mathsf{z}) = -\prod_{i=1}^{N} f_{M(x_i,\, p(\theta))}(y_i), \tag{27}$$

where $f_{M(x_i,p)}(y_i)$ is the PDF of the predicted output evaluated at $y = y_i$. This density function depends on f_p, which in turn, depends on θ. This performance metric is proportional to the probability of M reproducing the data sequence when $f_{p(\theta)}(p)$ is the joint PDF of p and the observations in z are IID. Note that (27) is a likelihood function where θ, not p, is the independent variable. The sign in (27) enables maximizing the likelihood via minimization. The calculation of (27) requires constructing an accurate approximation of $f_y(y, x_i)$ for all x_i's in z, a task that is more computationally expensive than the calculation of (26). The performance metric (26) is well suited to describe unimodal random processes. For instance, if $\mathsf{z} = \{(x_1, y_1), (x_1, y_2)\}$, the minimization of (26) will tend to yield a symmetric unimodal PDF for $y(x_1)$ centered about $(y_1 + y_2)/2$. Conversely, the metric in (27) is better suited to capture multi-modal processes. For instance, if $\mathsf{z} = \{(x_1, y_1), (x_1, y_2)\}$, the minimization of (27) will tend to yield a bimodal and symmetric PDF for $y(x_1)$ peaking at y_1 and y_2. The modality of the predicted random variable y is exclusively dependent on the structure of M and the spread of the observations. The framework above enables us posing an optimization program for calculating RPMs.

Optimization Program 3. *Consider the RPM defined by (18-20) for a norm d prescribed by any of the Equations (21-23) and any of the performance metrics in (26-27). Denote by*

Θ *the set of* θ *values leading to a feasible* P*. The value of the parameter* θ *required to fully prescribe this RPM*[5] *is given by*

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left\{ J(\theta, f_p, \mathsf{z}) : \underline{y}(x_i, P(\theta)) \leq y_i \leq \overline{y}(x_i, P(\theta)),\ 1 \leq i \leq N \right\}. \tag{28}$$

Therefore, this formulation targets the support set of the PDF in (19) such that (i) the resulting random process $R_y(x)$ maximizes the performance of the prediction, e.g., it minimizes the dispersion of probability about the observations, and (ii) all the observations in z can be reproduced by the computational model $M(x, p)$ for at least one parameter point of P. Condition (ii), which is enforced via the set of inequality constraints in (28), is equivalent to $P(\theta) \cap S_i \neq \emptyset$ for all $i = 1, \ldots, N$. Either of the parameterizations of P in (9) and (15) can be used in (28). In the context of (9) and (15), the constraint $\theta \in \Theta$ leads to $\alpha > 0$, and to $p_{\min} < p_{\max}$ respectively.

RPM Example

Next we calculate a RPM based on the same data sequence z^1 and the same computational model $M^1(x^1, p)$ used in the previously. Figure 6 shows the figure of merit $\gamma(p)$ defined by (20) and (23) for $N = 1$ (top-left), $N = 2$ (top-right), $N = 3$ (bottom-left), and $N = 50$ experimental observations (bottom-right). The top-left plot shows that the manifold S_1, which is the set of parameter points where $\gamma(p)$ takes on the maximum value, is a nonlinear function of p. The rate of decay in the value of $\gamma(p)$, thus of $f_p(p)$, is driven by the rate at which the function $(y_1^1 - M^1(p, x_1^1))^2$ departs from its zero manifold. This rate can be adjusted via parameters β_1 and β_2. The top-right figure is calculated by adding an additional term to the figure of merit shown in the top-left figure. This yields to a function $\gamma(p)$ where both manifolds S_1 and S_2 are superimposed. The bottom-left figure shows the figure of merit corresponding to S_1, S_2 and S_3. Note that $\gamma(p)$ reaches its largest values where the zero-prediction-error manifolds intersect. The bottom-right plot shows the figure of merit corresponding to all 50 manifolds. Note that two-regions of high-manifold concentration emerge, with the one in the vicinity of $p = [1, 1]^\top$ being dominant.

Figure 7 illustrates the dependency of the figure of merit $\gamma(p)$ on the tuning parameter β_2 for $N = 50$. Values of β_2 equal to one (top-left), 0.1 (top-right), 0.01 (bottom-right), and 0.001 (bottom-left) were used. Whereas larger values of β_2 yield faster rates of decay and therefore, more accurate probability allocations (i.e., allocations that concentrate more probability closer to the manifolds), smaller values yield a smother $\gamma(p)$, thus, density functions $f_p(p)$ that can be sampled more efficiently. Figures 6 and 7 are chosen to illustrate the β_2 dependency of $\gamma(p)$ and are not illustrative of an optimal support set for p.

A RPM defined by Equations (18-20, 23, 26), with support set P given by (15) and (28) was calculated. The prediction corresponding to the resulting RPM, to be called RPM2, is shown Figure 8. This formulation aims at concentrating the probability of the predicted response as close as possible to the observations. The comparison of Figures 9 and 8 illustrates

[5]The parameters β_1 and β_2 in (20) used to prescribe $f_p(p)$ could be additional optimization variables.

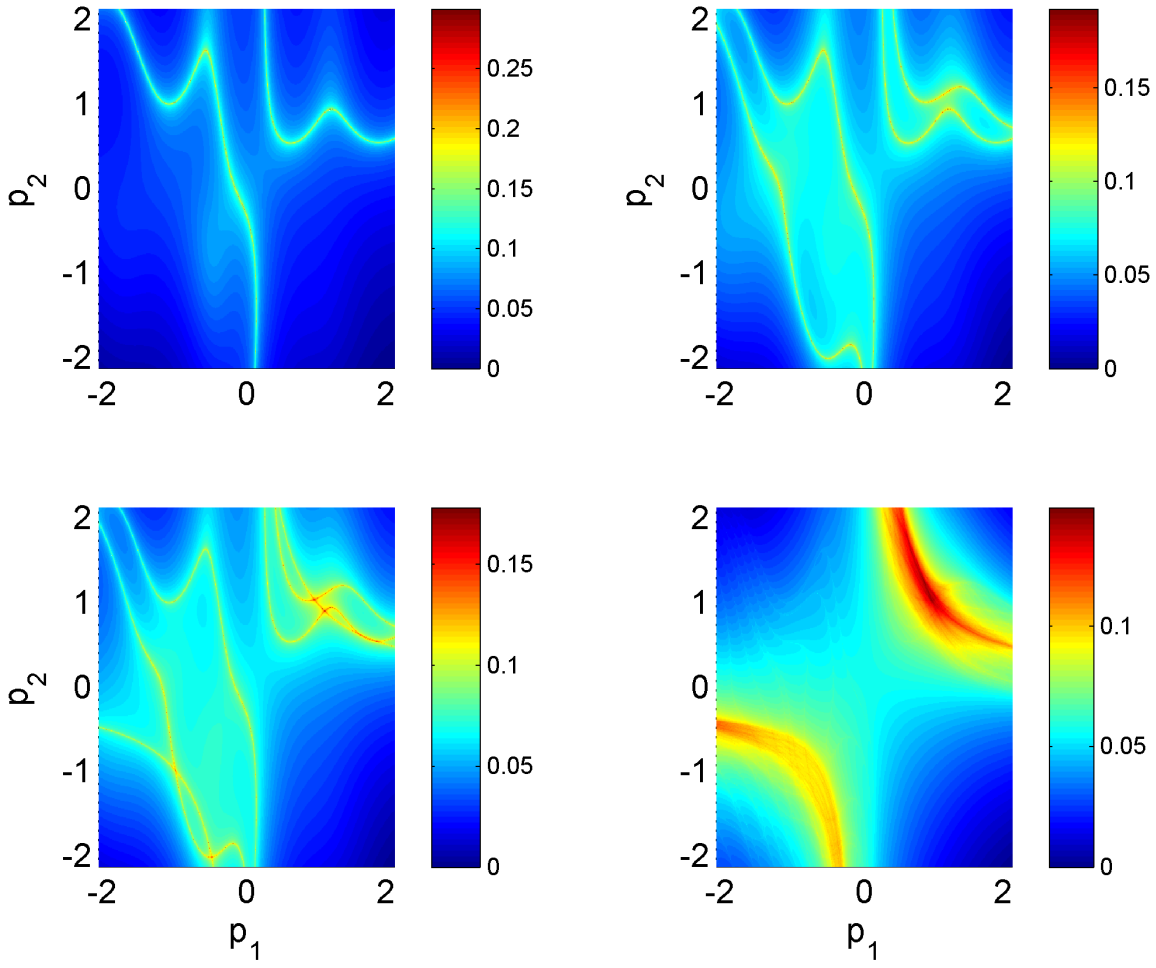

Figure 6. Dependence of $\gamma(p)$ on the number of observations N.

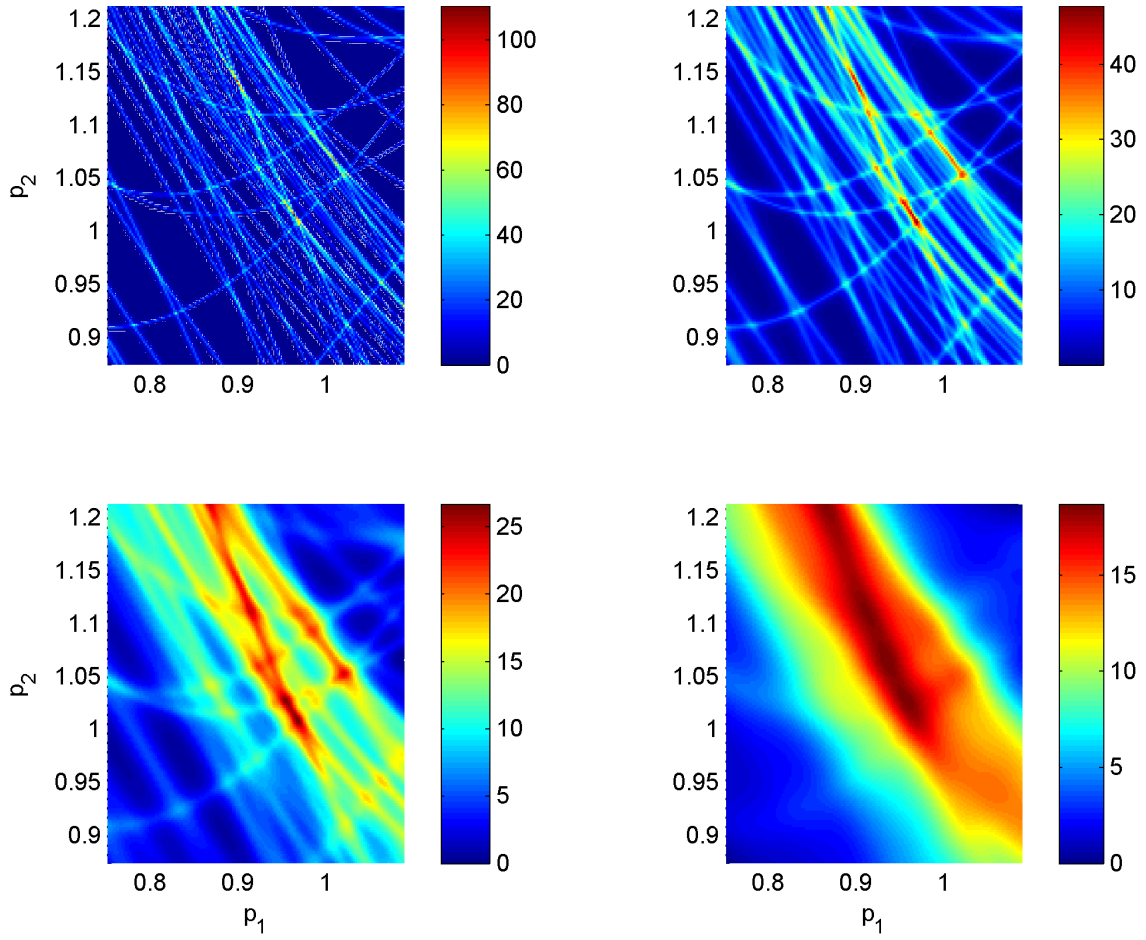

Figure 7. Dependence of $\gamma(p)$ on the tuning parameter β_2.

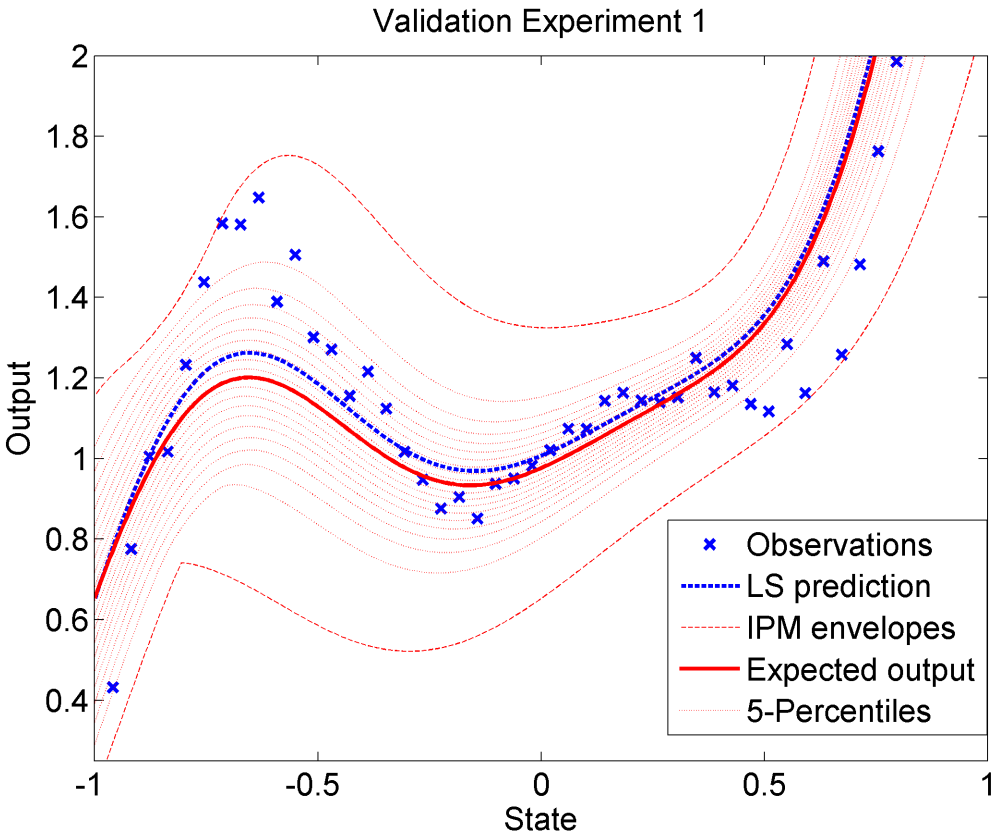

Figure 8. RPM with minimal dispersion about observations (RPM2).

that the mean response and the 5-percentile curves corresponding to RPM2 are a tighter representation of the data. Note that the mean prediction and the LS prediction are close even though they have a different mathematical structure (i.e., one corresponds to the evaluation of the computational model at a single parameter point whereas the other one is the average value of the model predictions corresponding to a large number of parameter points). The support set of p corresponding to RPM2, $P(\hat{\theta}_1)$, is bounded by $\hat{p}_{\min} = [0.8030,\ 0.8106]^\top$ and $\hat{p}_{\max} = [1.0356,\ 1.2247]^\top$, for which $J(\hat{\theta}, f_p, \mathsf{z}^1) = 3.8140$. This region contains but it is not centered about points where $f_p(p)$ takes on its largest values. Conversely, the performance of RPM1 is $J(\theta, f_p, \mathsf{z}^1) = 8.1332$. This poor performance is expected since the probabilistic features of RPM1 are not optimal.

For comparison purposes, model $M^1(x^1, p)$ was calibrated using Bayesian inference. The calibration is based in Equation (1) where η is a zero-mean Gaussian with fixed variance σ^2. Calibration yields a probabilistic description of $[p, \sigma]$. The resulting RPM is shown in Figure 9, where the maximum likelihood prediction $M^1(x^1, p_{\text{MLE}})$ for the maximum likelihood estimate (MLE) $p_{\text{MLE}} = [1.8071,\ 0.8666]^\top$ is also shown. Note that the MLE prediction and

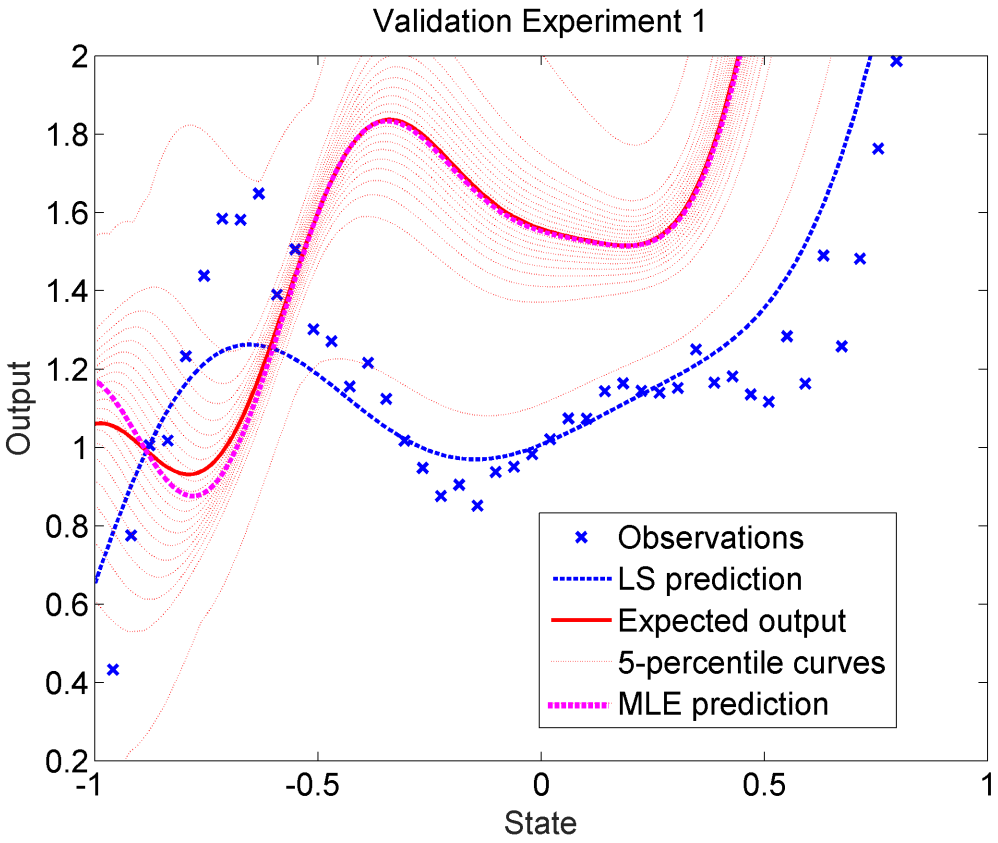

Figure 9. RPM resulting from Bayesian Inference.

the random process are poor predictors of both the spread and overall trend of the data.

Uncertainty Roll-up

Uncertainty quantification aims at studying the effects of uncertainty in full-system target applications. Unfortunately, we often have some tests at the subsystem/component level but not at the full system level. We want to use these tests to calibrate/validate the computational models of the subsystems in order to generate informative predictions of the target application. By roll-up we mean the process of calibrating the parameters of one or several computational models and using the resulting characterization of p to make predictions of a target application which also depends on p. In particular, we want to prescribe p according to the ability of the computational model $M^j(x^j, p)$ to reproduce the data sequence z^j for all $j = 1, \ldots, n_v$. Note that the structure of the DGM and of the computational model corresponding to a validation experiment might be different from those for other validation experiments. Once p is characterized, a prediction for the target application $y^{\mathrm{t}} = M^{\mathrm{t}}(x^{\mathrm{t}}, p)$

will be made. Note that the response quantities y^1, y^2, ..., y^{n_v} and y^{t}; as well as the corresponding states x^1, x^2, ..., x^{n_v} and x^{t} may differ from each other (they might describe different physics and different quantities of interest).

The IPM and RPM formulations above can be naturally extended to study the roll-up problem. The main extension considered in this section is the joint usage of all the data sequences with all computational models. The resulting characterization of p along with each computational model M^i constitutes a predictor model. Therefore, we will obtain a family of predictor models linked by the common prescription of the uncertainty in p where, each family member describes one of the n_v validation experiments. Extensions to the formulations above are presented next.

Optimization Program 4. *Consider the family of IPMs defined by (3-9) for $M^1(x^1, p)$, $M^2(x^2, p)$, ..., $M^{n_v}(x^{n_v}, p)$ and assume that the values c and $m > 0$ are set in advance. The value of the parameter α required to fully prescribe this family is given by*

$$\hat{\alpha} = \underset{\alpha>0}{\operatorname{argmin}} \left\{ \frac{1}{n_v} \sum_{j=1}^{n_v} \frac{E_{x^j}[\delta_{y^j}(x^j, P(\theta))]}{D^j} : \ \underline{y}^j(x^j_{i^j}, P(\theta)) \leq y^j_{i^j} \leq \overline{y}^j(x^j_{i^j}, P(\theta)), \right. \tag{29}$$
$$\left. 1 \leq j \leq n_v, \ \ 1 \leq i^j \leq N^j, \ \text{where } \theta = \{c, m, \alpha\} \ \right\},$$

and D^j is a normalization constant.

Therefore, (29) yields a P that minimizes the average spread of the n_v predicted responses, such that the envelopes of all such responses contain all observations. The normalization constant D_j is used to make the spread of different response metrics comparable. One possible choice for this constant is the variance of the observed responses associated with the same validation experiment $D^j = V_i[y^j_i]$. As expected, the spread of the prediction envelopes for all n_v models, as well as the size of P, grow with the value of n_v. The extension of Optimization Program 2, which follows the same structure and rationale of (29), can be easily inferred. Regarding RPMs, the extension of Optimization Program 3 is as follows.

Optimization Program 5. *Consider the family of RPMs defined by (18-20) for $M^1(x^1, p)$, $M^2(x^2, p)$, ..., $M^{n_v}(x^{n_v}, p)$, and a norm d prescribed in any of the Equations (21-23). The value of the parameter θ required to fully prescribe this family is given by*

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left\{ \bar{J}(\theta, f_p, \mathsf{z}^1, \ldots, \mathsf{z}^{n_v}) : \ \underline{y}^j(x^j_{i^j}, P(\theta)) \leq y^j_{i^j} \leq \overline{y}^j(x^j_{i^j}, P(\theta)), \right.$$
$$\left. 1 \leq j \leq n_v, \ \ 1 \leq i^j \leq N^j \right\}, \tag{30}$$

where the cost function $\bar{J}$ is given by

$$\bar{J}(\theta, f_p, \mathsf{z}^1, \ldots, \mathsf{z}^{n_v}) = \sum_{j=1}^{n_v} \frac{q^j J(\theta, f_p, \mathsf{z}^j)}{D^j}, \tag{31}$$

$$\bar{J}(\theta, f_p, \mathsf{z}^1, \ldots, \mathsf{z}^{n_v}) = -\prod_{j=1}^{n_v} \left| J(\theta, f_p, \mathsf{z}^j) \right|^{\eta_j}, \tag{32}$$

when the performance metric J is given by (26) or (27) respectively. The parameters $q \in \mathbb{R}^{n_v}$ and $\eta \in \mathbb{R}^{n_v}$ are constants set by the analyst, while D^j is the normalization constant introduced earlier.

Therefore, (30) yields a value of θ for which the corresponding $f_p(p)$ minimizes the cost function $\bar{J}$ such that the envelopes of the predicted responses contain all corresponding observations. The cost $\bar{J}$ in Equation (31) is the weighted sum of the dispersion of all RPMs about the observations. Use $q = 1$ to give the same importance to the calibration of all RPMs. Conversely, the cost $\bar{J}$ in Equation (32) depends on the likelihood of all RPMs reproducing the corresponding data sequence. Use $\eta_j = 1$ to weigh each observation point equally, and $\eta_j = 1/N^j$ to weigh each data sequence equally. The spread of the envelopes and the tightness of the random predictions resulting from (30), regardless of the particular performance function used, can only degrade as the value of n_v increases. As before, either of the parameterizations of P in (9) and (15) can be used in (30). In the context of (9) and (15), the constraint $\theta \in \Theta$ leads to $\alpha > 0$, and to $p_{\min} < p_{\max}$ respectively.

Roll-up Example

We now consider *Validation Experiment 2*. The DGM associated to this experiment, from which we obtained the data sequence z^2 with $N^2 = 50$ observations, depends nonlinearly on the state and parameters, and it is subject to state dependent random noise. Figure 10 shows the observations. The computational model $M^2(p, x^2)$, for which $x^2 \in \mathbb{R}$ and $p \in \mathbb{R}^2$, was built to describe this DGM. As compared to the DGM, the model M^2 is subject to epistemic-, aleatory- and model-form uncertainty.

We will calculate a RPM based on z^2 and $M^2(p, x^2)$ first. This RPM is defined by (18-20) with norm d given by (23) and performance metric given by (26). The corresponding LS parameter estimate, $p_{\text{LS2}} = [1.3090, 1.0866]^\top$, differs significantly from the LS parameter estimate for Validation Experiment 1, $p_{\text{LS1}} = [0.9420, 1.0689]^\top$ (see Fixed-parameters IPM Example). This difference is caused by the uncertainty affecting the models and the noise affecting the DGMs. The resulting RPM, to be called RPM3, is shown in Figure 10. The performance of RPM3, $J(\theta, f_p, \mathsf{z}^2) = 77.6056$, is much larger than that of RPM2. The support set of p corresponding to RPM3, $P(\hat{\theta}_2)$, is given by $\hat{p}_{\min} = [1.1333,\ 0.3192]^\top$ and $\hat{p}_{\max} = [1.5174,\ 1.8015]^\top$. Note that the uncertainty affecting M^2 is significantly larger than that of M^1.

The simultaneous calibration of M^1 and M^2 using z^1 and z^2 is considered next. The mathematical structure of the DGMs associated to each validation experiment, thus of the data sequences, and of the corresponding computational models is significantly different. RPMs for both validation experiments were calculated using Equations (23, 26, 30) for $\eta = 1$. These RPMs, to be called RPM4 and RPM5, are shown in Figures 11 and 12. These figures show the 5-percentile curves, the IPM envelopes, and the LS predictions corresponding to each validation experiment alone. The large differences between p_{LS1} and p_{LS2}, and between the corresponding predictions illustrate that the uncertainty for the joint system is much

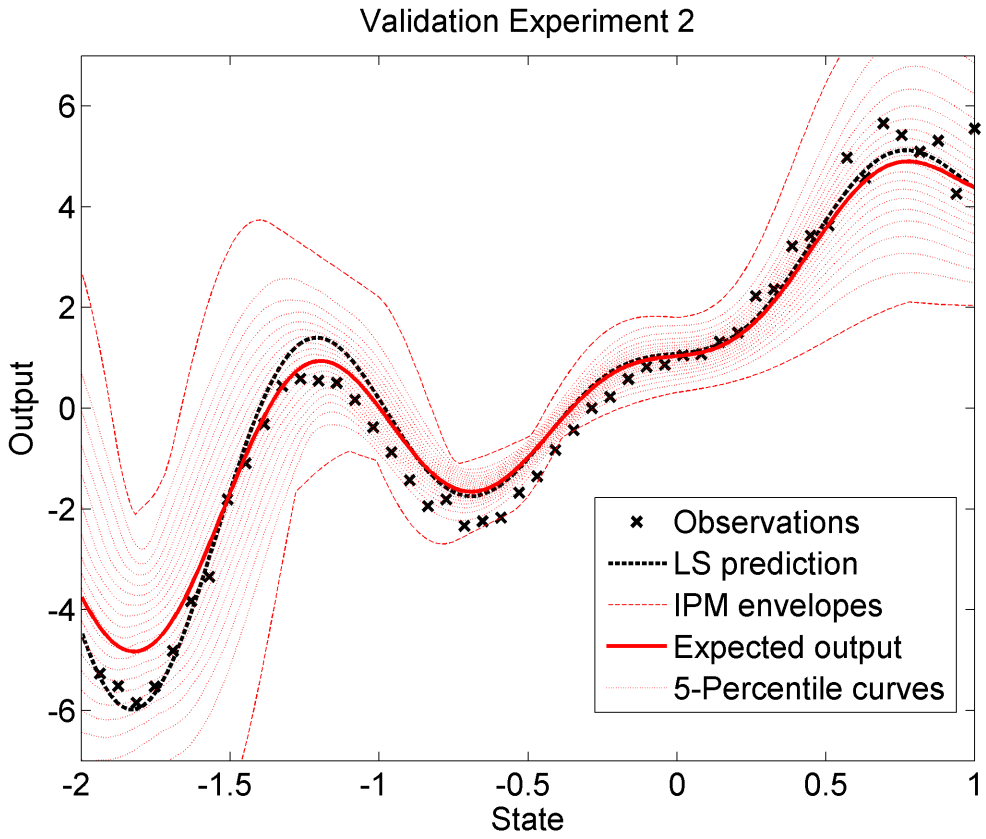

Figure 10. RPM for Validation Experiment 2 (RPM3).

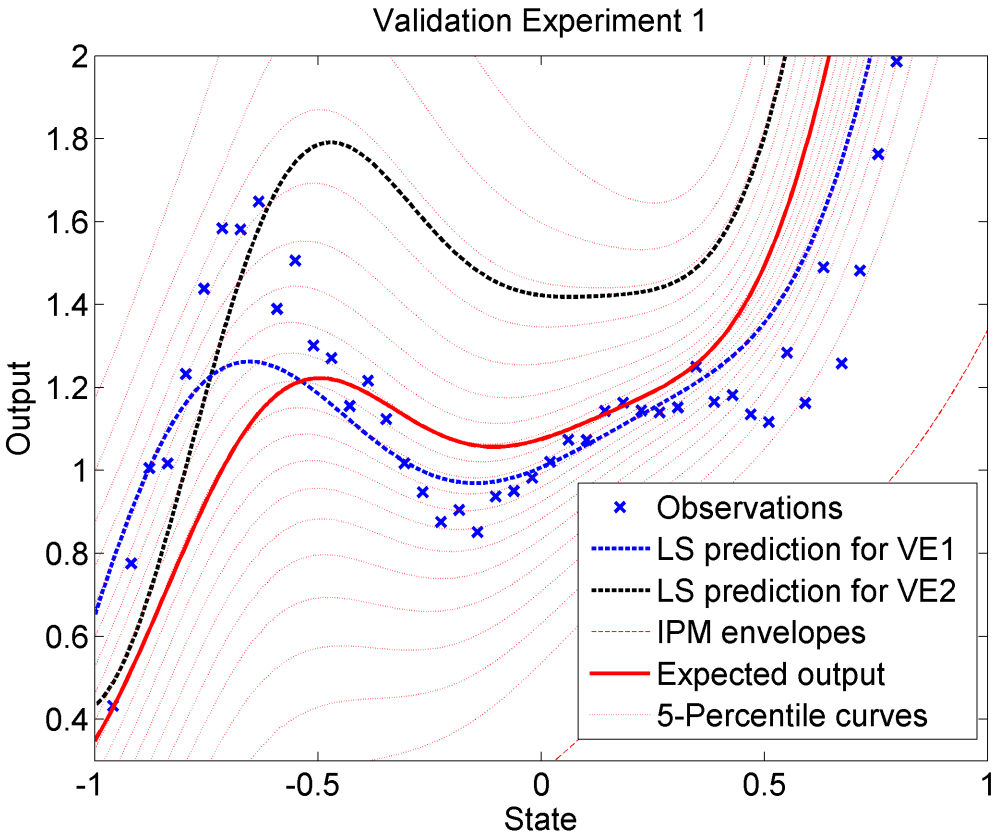

Figure 11. RPM for Validation Experiment 1 (RPM4).

larger than the uncertainty resulting from calibrating each computational model individually. The support set of both RPM4 and RPM5, $P(\hat{\theta}_3)$, is bounded by $\hat{p}_{\min} = [0.8049,\ 0.3413]^\top$ and $\hat{p}_{\max} = [1.5173,\ 1.8018]^\top$. Note that $P(\hat{\theta}_3) \neq P(\hat{\theta}_1) \cup P(\hat{\theta}_2)$, and the volume of $P(\hat{\theta}_3)$ is much larger than the other two.

The comparison of Figures 8 and 11 indicates a considerable degradation in the quality of the prediction. The performance of RPM4, $J(\theta, f_p, \mathsf{z}^1) = 30.2080$, is 7.92 times larger than RPM2's; whereas the performance of RPM5, $J(\theta, f_p, \mathsf{z}^2) = 211.4523$, is about 2.72 time larger than RPM3's. The degradation in the tightness of the prediction is caused by calibrating M^1 and M^2 jointly.

Figure 13 provides insight into the process by which the probabilistic description of the uncertainty p is constructed. This figure shows the joint PDFs corresponding to the RPMs resulting from calibrating M^1 alone (left-subplot), from calibrating M^2 alone (center-subplot), and from calibrating M^1 and M^2 together (right-subplot). The concentration of probability indicates how the zero-prediction error manifolds are distributed in the parameter space.

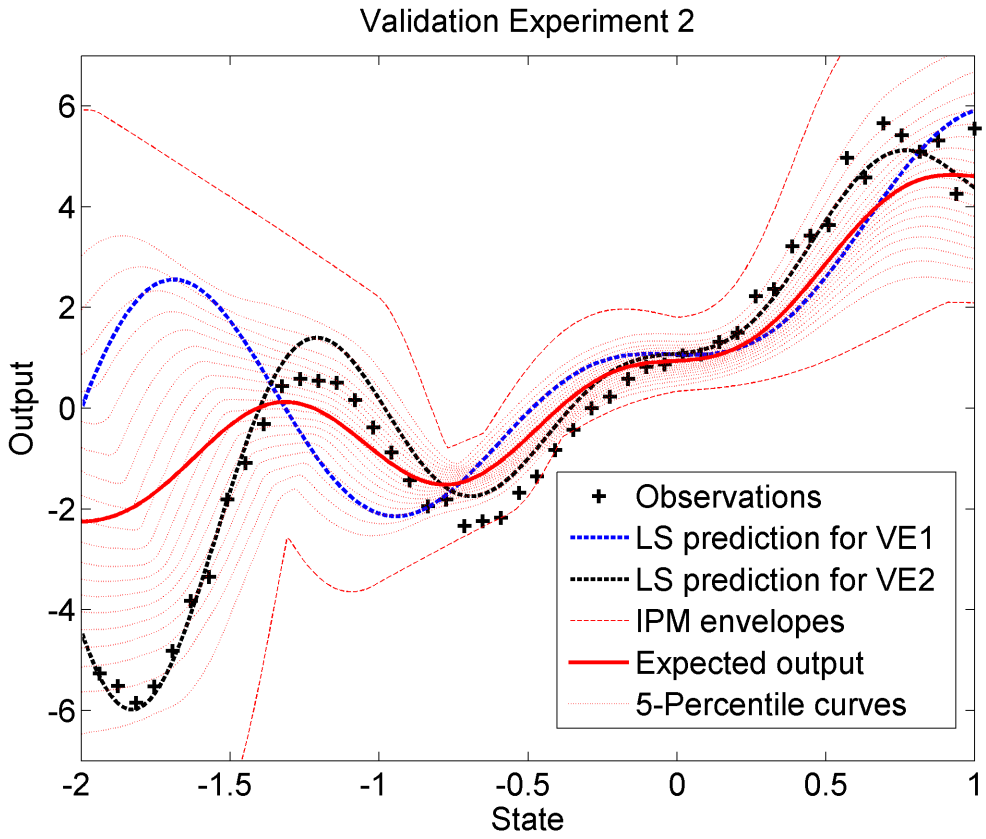

Figure 12. RPM for Validation Experiment 2 (RPM5).

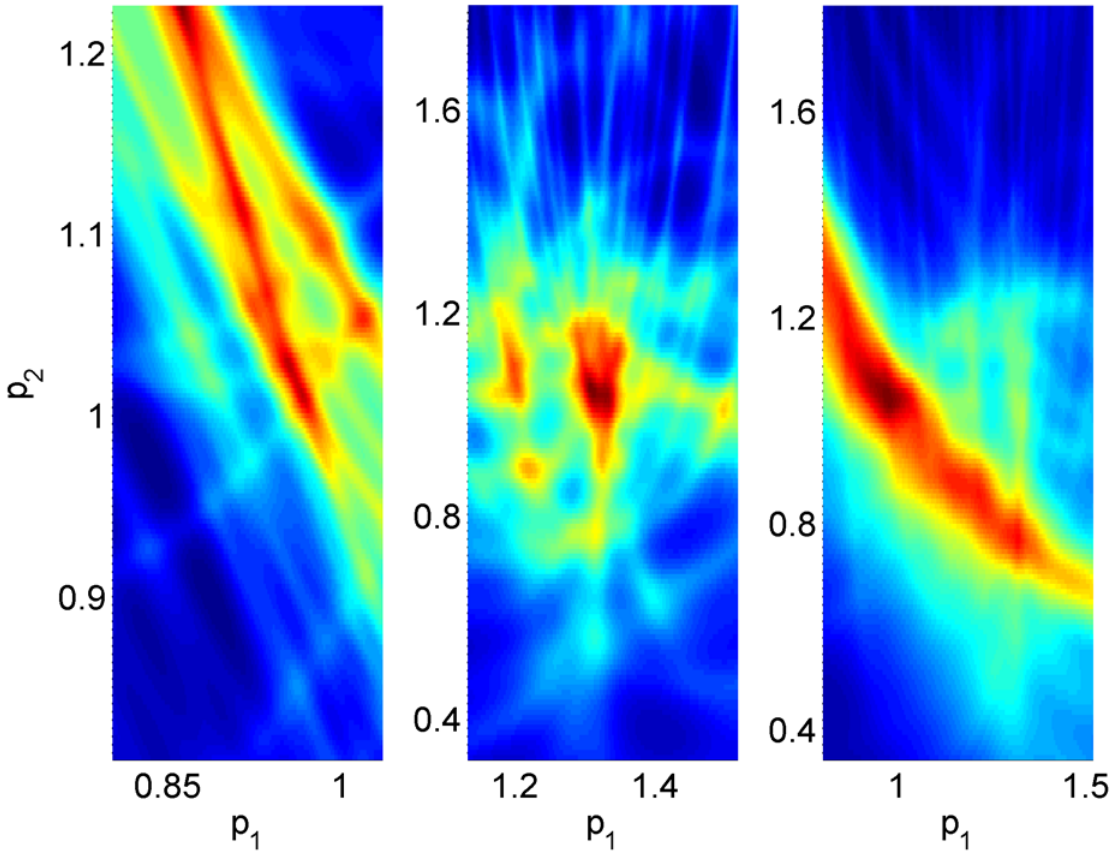

Figure 13. PDFs of RPM2 (left), RPM3 (center), and RPM4-RPM5 (right).

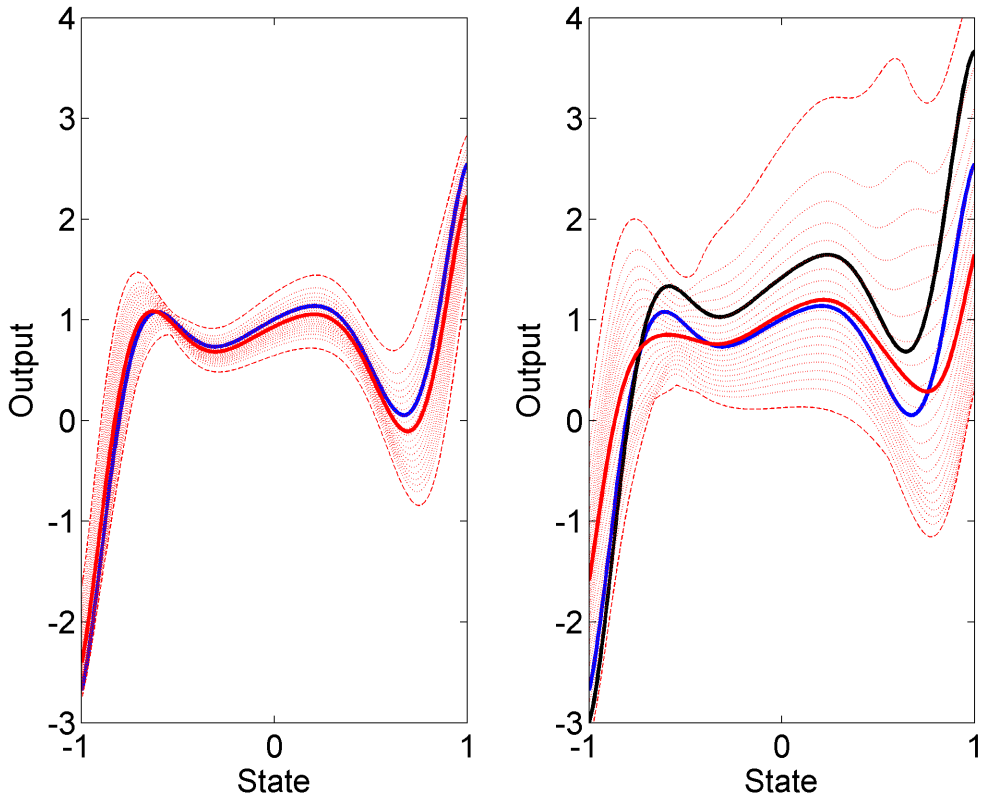

Figure 14. RPM of target application for calibrations based on 1st (left) and both (right) validation experiments. Same line conventions used previously apply.

Finally, we use the characterizations of p obtained above to make predictions on a target application. Figure 14 shows the RPM for the target application $y^{\mathrm{t}} = M^{\mathrm{t}}(x^{\mathrm{t}}, p)$ when the uncertainty in p is characterized based on the first validation experiment (left subplot), and on the first and second validation experiments (right subplot). The LS predictions corresponding to the first (blue), and second (black) validation experiments, the 5-percentile curves (red-dotted), process support (red-dashed) and the expected output (red-solid) are shown. The degradation in the predictive response, in terms of the spread of the support and of the dispersion of probability, is apparent. Note that the predicted random process is skewed towards smaller output values. Further notice that whereas the spread of the output increases when both validation experiments are used for calibration, the variation of the mean output response has actually decreased. Notice that the prediction for the target application depends on the ability of M^1 to reproduce the sequence z^1, on the ability of M^2 to reproduce the sequence z^2, and on the structure of M^{t}.

Conclusions

This paper proposes a new paradigm for the calibration of computational models and a framework for uncertainty roll-up. The formulations proposed render a description of the uncertainty in the models parameters, thereby enabling the evaluation of the dispersion in model's predictions. The process by which the predictor models are generated is based on modeling the discrepancy between model predictions and observations using deterministic and probabilistic means. In contrast to standard Bayesian approaches, the predictions resulting from IPMs and RPMs wont converge to a deterministic response function as more data is made available. Instead, IPMs converge to an interval valued function matching the spread of the experimental data, whereas RPMs converge to a random process matching their distribution.

References

[1] J. P. Aubin and A. Cellina. *Differential Inclusions*. Springer-Verlag, Berlin, Germany, 1984.

[2] J. P. Aubin, J. Lygeros, M. Quincampoix, S. Sastry, and N. Seube. Impulse differential inclusions: a viability approach to hybrid systems. *IEEE Transactions on Automatic Control*, 47(1):2–20, 2002.

[3] L. G. Crespo, S. P. Kenny, and D. P. Giesy. Interval predictor models with a formal characterization of uncertainty and reliability. In *53rd IEEE Conference on Decision and Controls*, December 15-17, Los Angeles, CA, USA 2014.

[4] M.C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society B*, 63(3):425–464, 2001.

[5] T. P. Lane and W. H. DuMouchel. Simultaneous confidence intervals in multiple regression. *The American Statistician*, 48(4):315–321, 1994.

[6] M. Milanese and C. Novara. Set-membership identification of nonlinear systems. *Automatica*, 40(6):957–975, 2004.

[7] M. Milanese and C. Novara. Set-membership prediction of nonlinear time systems. *IEEE Transactions on Automatic Control*, 50(11):1655–1669, 2005.

[8] G. A. Seber and C. J. Wild. *Nonlinear Regression*. JohnWiley & Sons, Hoboken, New Jersey, USA, 2003.

DISTRIBUTION:

1 MS 0899 Technical Library, 9536 (electronic copy)

Sandia National Laboratories